

Core Research

Our core research efforts will continue to emphasize basic research on AI techniques applicable to biomedical problems and the generalization and documentation of tools to facilitate and broaden application areas.

SUMEX core research funding is complementary to similar funding from other agencies and contributes to the long-standing interdisciplinary effort at Stanford in basic AI research and expert system design. We expect this work to provide the underpinnings for increasingly effective consultative programs in medicine and for more practical adaptations of this work within emerging microelectronic technologies. Specific aims include:

- 1) Continue to explore basic artificial intelligence issues for knowledge acquisition, representation, and utilization; reasoning in the presence of uncertainty; strategy planning; and explanations of reasoning pathways with particular emphasis on biomedical applications.
- 2) Support community efforts to organize and generalize AI tools that have been developed in the context of individual application projects. This will include work to organize the present state-of-the-art in AI techniques through the AI Handbook effort and the development of practical software packages (e.g., AGE, EMYCIN, UNITS, and EXPERT) for the acquisition, representation, and utilization of knowledge in AI programs. The objective is to evolve a body of software tools that can be used to more efficaciously build future knowledge-based systems and explore other biomedical AI applications.

Hardware Acquisition Rationale

As discussed in our progress report and supported by collaborating project reports, we have implemented an effective set of computing resources to support AI applications to biomedical research. At the resource core is the KI-TENEX/2020 facility, augmented by portions of the Rutgers 2050 and Stanford SCORE 2060 machines. These have provided an unsurpassed set of tools for the initial phases of SUMEX-AIM development in terms of operating system facilities, human engineering, language support for artificial intelligence program development, and community communications tools. As the size of our community and the complexity of knowledge-based programs have increased, several issues have become important for the continued development and practical dissemination of AI programs:

- 1) The community has a continuing need for more computing capacity. This arises from the growth of new applications projects, new core research ideas, and the need to disseminate mature systems within and outside of the AIM community. Nowhere is this felt more strongly than among the Stanford community where system access constraints have seriously impeded development progress. A picture of system congestion can be found in the summary of loading statistics beginning on page 29 and in the statements from many of our user projects.
- 2) Many programs require a larger virtual address space. As AI systems become more expert and encompass larger and more complex domains, they require ever larger knowledge bases and data structures that must be traversed in the course of solving problems. The 256K word address limit of the PDP-10 has constrained program development as discussed in our renewal proposal. Increasing effort has gone into "overlays" resulting in higher machine overhead, more difficulty in making program changes, and lost programmer time. Simpler hardware solutions are needed.
- 3) AI programs are being tested and disseminated increasingly beyond their development communities. We cannot continue to provide all of the computing resources this implies through central systems like SUMEX. The capacity does not exist. Network communications facilities are not able to support facile human interactions (high speed, improved displays, graphics, and speech/touch modalities). And a grant-supported research environment cannot meet the technical and administrative needs of a "production" community. Thus, we need to explore better ways to package complex AI software and distribute the necessary computing tools cost effectively into the user communities.

No single solution to these requirements for future development is available and we proposed and got peer approval to investigate a variety of machine architectures and support functions over the next grant period including:

- 1) experimentation with new shared centralized systems
- 2) distributed single-user "professional workstations"
- 3) improved communications tools to integrate them together effectively.

In addition to continuing operation of the existing resources, we plan to direct SUMEX research efforts to explore the potential of such newly available systems as solutions to AIM community needs. Our approach will be to integrate a heterogeneous set of network-connected hardware tools, some of which will be distributed through the user community. We will emphasize the development of system and application level software tools to allow effective use of these resources and continue to provide community leadership to encourage scientific communications.

Specific Hardware Plans for Year 09

In our proposal as approved by council, we described a carefully detailed plan for hardware acquisition. One of the approved purchases, the augmentation to the AMPEX core memory for the KI-10 duplex, was approved for the current year 08 and has already been implemented. In addition, for the technical reasons discussed on page 11, we have obtained BRP approval to accelerate the purchase of the five approved professional workstations to year 09 and to delay the first VAX purchase to year 10. The following then is a summary of planned hardware purchase for year 09:

- Buy five Interlisp Dolphin professional workstations for use in developing and experimenting with this means for AI program export and human interface enhancements.
- Develop a file server coupled to SUMEX host machines via the high speed Ethernet. This will minimize the need for redundant large file systems on each host and alleviate the file storage limitations of the AIM community.
- Acquire examples of state-of-the art display equipment including a bit-mapped display station and a hardcopy laser printing device.
- Buy additional required communications, interface, and test equipment to support the above acquisitions and community needs.

Continued Operation of Existing Hardware

The current SUMEX-AIM facilities represent a large existing investment. We do not propose any substantial changes to the existing KI-10 and 2020 hardware systems and we expect them to continue to provide effective community support and serve as a communication nucleus for more distributed resources. The proposed augmentation of the existing KI-10 AMPEX memory box in order to reduce page swapping overhead is underway.

It should be recognized that the KI-10 processors are now 6 years old and will be 12 years old at the end of the proposed grant term. We have already begun to feel maintenance problems from age such as poor electrical contacts from oxidization and dirt, backplane insulation flowing on "tight wraps", and brittle cables. These problems are quite manageable still and we expect to be able to continue reliable operation over the next grant term.

We plan no upgrades to the 2020 configuration. The current file shortage will be remedied in conjunction with that of the rest of the facility by implementing a community file server sharable and accessible via the Ethernet.

For both systems, we are actively working to complete efficient interfaces to the Ethernet to allow flexible, high speed terminal connections, file transfers, and effective sharing of network, printing, plotting, remote links, and other resources. This system will form the backbone for smooth integration of future hardware additions to the resource.

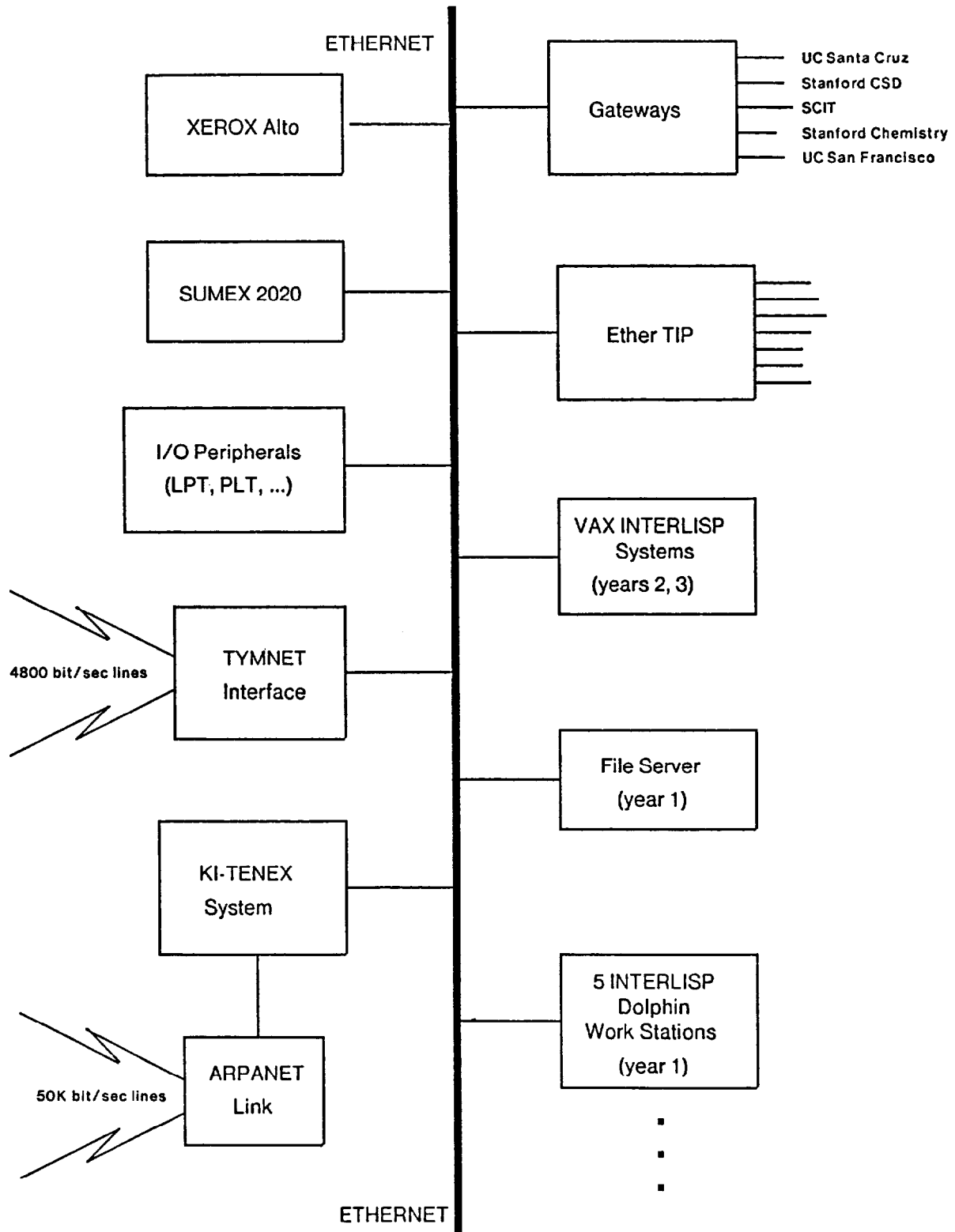


Figure 14. Planned Ethernet System to Integrate System Hardware

Communication Networks

Networks have been centrally important to the research goals of SUMEX-AIM and will become more so in the context of increasingly distributed computing. Communication will be crucial to maintain community scientific contacts, to facilitate shared system and software maintenance based on regional expertise, to allow necessary information flow and access at all levels, and to meet the technical requirements of shared equipment.

Long-Distance Connections

We have had reasonable success at meeting the geographical needs of the community during the early phases of SUMEX-AIM through our ARPANET and TYMNET connections. These have allowed users from many locations within the United States and abroad to gain terminal access to the AIM resources (SUMEX, Rutgers, and SCORE) and through ARPANET links to communicate much more voluminous file information. Since many of our users do not have ARPANET access privileges for technical or administrative reasons, a key problem impeding remote use has been the limited communications facilities (speed, file transfer, and terminal handling) offered currently by commercial networks. Commercial improvements are slow in coming but may be expected to solve the file transfer problem in the next few years. A number of vendors (AT&T, IBM, Xerox, etc.) have yet to announce commercially available facilities but TELENET is actively working in this direction. We plan to continue experimenting with improved facilities as offered by commercial or government sources in the next grant term. We have budgeted for continued TYMNET service and an additional amount annually for experimental network connections.

High-speed interactive terminal support will continue to be a problem since one cannot expect to serve 1200-9600 baud terminals effectively over shared long-distance trunk lines with gross capacities of only 9600-19200 baud. We feel this is a problem that is best solved by distributed machines able to effectively support terminal interactions locally and coupled to other AIM machines and facilities through network or telephonic links. As new machine resources are introduced into the community, we will allocate budgeted funds with Executive Committee advice to assure effective communication links.

Local Intermachine Connections

A key feature of our plans for future computing facilities is the support of a heterogeneous processing environment that takes advantage of newly available technology and shared equipment resources between these machines. The "glue" that links these systems together is a high speed local network. We have chosen Ethernet and the Xerox PUP [9, 12] protocols for these interconnections. This choice was based on the availability of that technology now and the economics of using already developed TENEX and other server software. We expect the Ethernet system to continue to meet our technical needs for the coming grant term and we plan to continue to use it. We are working closely with other groups here

at Stanford and elsewhere to share hardware interface and software designs wherever possible.

Our goals are to complete integration of the SUMEX-AIM system, including making selected KI-10 peripherals available as Ethernet nodes, creating links to nearby campus resources, and establishing needed remote links to other groups not on the ARPANET such as Wipke at the University of California at Santa Cruz. A diagram of our Ethernet system is shown in Figure 14 on page 55 and includes the following major elements:

- 1) KI-10 direct memory access interface. We currently have an inefficient I/O bus connection.
- 2) Stanford campus gateway. Establish links to other Ethernets on campus to allow access to special resources (Dover printer, plotters, typesetting equipment, etc.) and to allow users to easily access various computing resources.
- 3) Ethertip. We need additional terminal ports into the system and the Ethernet provides a natural mechanism to do this supporting high speed terminals and connections to various resources (KI-10, 2020, VAX's, etc.).
- 4) TYMNET connection. This connection currently comes through the KI-10's and will be moved to a separate Ethernet node. This will free the KI-10's from handling the special TYMNET protocol and will allow TYMNET users to access any of the SUMEX-AIM resources. Similar facilities for the ARPANET may also be implemented depending on administrative constraints.
- 5) Printer/plotter service. We plan to make these local resources accessible from any of the SUMEX-AIM machines instead of being centered on the KI-10's. This will also free up the KI-10's from routine spooler tasks.
- 7) Connections for other machines (VAX's, Professional Workstations, file server, etc.)

Resource Software

We will continue to maintain the existing system, language, and utility support software on our systems at the most current release levels, including up-to-date documentation. We will also be extending the facilities available to users where appropriate, drawing upon other community developments where possible. We rely heavily on the needs of the user community to direct system software development efforts. Specific development areas for existing systems include:

- 1) completion of the Ethernet connections and necessary host software. This will include basic packet handling, PUP protocols at all levels, and relocation of shared existing resources to become Ethernet nodes.
- 2) bug fixes in the current monitors. We still have a number of bugs that cause infrequent crashes and that are hard to isolate because they cause system problems long after the fact. We will continue to work to repair these problems as time permits.
- 3) continued evaluation of system efficiency to improve performance.
- 4) compatibility issues. Our current compatibility package for TOPS-20 requires additional work to extend its features. We will also keep it up-to-date as DEC make new changes to their system.
- 5) continued work to create similar working and programming environments between our TENEX and TOPS-20 systems. This will include moving TENEX features like the SUMEX GTJFN enhancements and scheduling controls as needed to TOPS-20 and vice versa
- 6) continued work to improve system information and help facilities for users.

Our plans for augmenting the SUMEX-AIM resources will entail substantial new system and subsystem programming. Our goals will be to derive as much software as possible from the user communities of the new VAX and Professional Workstation machines but we expect to have to do considerable work to adapt them to our biomedical AI needs. Many features of these systems are designed for a computer science environment and lack some of the human engineering and "friendliness" capabilities we have found needed to allow non-computer scientists to effectively use them. We are beginning to experiment with physician needs for interfaces to our AI programs to be better able to adapt the new machines as professional aids. Also many of the utility tools that we take for granted in the well-developed TENEX and TOPS-20 environment (communications, text manipulation, file management, accounting, etc.) will have to be reproduced. We expect to set up many of the common information services as network nodes.

Within the AIM community we expect to serve as a center for software sharing between various distributed computing nodes. This will include contributing locally developed programs, distributing those derived from

elsewhere in the community, maintaining up-to-date information on subsystems available, and assisting in software maintenance.

Community Management

We plan to retain the current management structure that has worked so well. We will continue to work closely with the management committees to recruit the additional high quality projects which can be accommodated and to evolve resource allocation policies which appropriately reflect assigned priorities and project needs. We expect the Executive and Advisory Committees to play an increasingly important role in advising on priorities for facility evolution and on-going community development planning in addition to their recruitment efforts. The composition of the Executive committee will grow as needed to assure representation of major user groups and medical and computer science applications areas. The Advisory Group membership rotates regularly and spans both medical and computer science research expertise. We expect to maintain this policy.

We will continue to make information available about the various projects both inside and outside of the community and thereby promote the kinds of exchanges exemplified earlier and made possible by network facilities.

The AIM workshops under the Rutgers resource have served a valuable function in bringing community members and prospective users together. We will continue to support this effort. This summer the AIM workshop will be held in Vancouver, British Columbia in conjunction with the International Joint Conference on Artificial Intelligence. We are actively helping to organize the meeting. We will continue to assist community participation and provide a computing base for workshop demonstrations and communications. We will also assist individual projects in organizing more specialized workshops as we have done for the DENDRAL and AGE projects.

We plan to continue indefinitely our present policy of non-monetary allocation control. We recognize, of course, that this accentuates our responsibility for the careful selection of projects with high scientific and community merit.

Training and Education Plans

We have an on-going commitment, within the constraints of our staff size, to provide effective user assistance, to maintain high quality documentation of the evolving software support on the SUMEX-AIM system, and to provide software help facilities such as the HELP and Bulletin Board systems. These latter aids are an effective way to assist resource users in staying informed about system and community developments and solving access problems. We plan to take an active role in encouraging the development and dissemination of community databases such as the AI Handbook, up-to-date bibliographic sources, and developing knowledge bases. Since much of our community is geographically remote from our machine, these on-line aids are indispensable for self help. We will continue to provide on-line personal assistance to users within the capacity of available staff through the SNDMSG and LINK facilities.

We budget funds to continue the "collaborative linkage" support initiated during the first term of the SUMEX-AIM grant. These funds are allocated under Executive Committee authorization for terminal and communications support to help get new users and pilot projects started.

Finally, we will continue to actively support the AIM workshop series in terms of planning assistance, participation in program presentations and discussions, and providing a computing base for AI program demonstrations and experimentation.

Core Research Plans

SUMEX core research includes both basic AI research and development of community tools useful for building expert systems. Expert systems are symbolic problem solving programs capable of expert-level performance, in which domain-specific knowledge is represented and used in an understandable line of reasoning. The programs can be used as problem solving assistants or tutors, but also serve as excellent vehicles for research on representation and control of diverse forms of knowledge. MYCIN is one of the best examples.

Because the main issues of building expert systems are coincident with general issues in AI, we appreciate the difficulty of proposing to "solve" basic problems. However, we do propose to build working programs that demonstrate the feasibility of our ideas within well defined limits. By investigating the nature of expert reasoning within computer programs, the process is "demystified". Ultimately, the construction of such programs becomes itself a well-understood technical craft.

The foundation of all of our core research work is expert knowledge: its acquisition from practitioners, its accommodation into the existing knowledge bases, its explanation, and its use to solve problems. Continued work on these topics provides new techniques and mechanisms for the design and construction of knowledge-based programs; experience gained from the actual construction of these systems then feeds back both (a) evaluative information on the ideas' utility and (b) reports of quite specific problems and the ways in which they have been overcome, which may suggest some more general method to be tried in other programs.

One of our long-range goals is to isolate AI techniques that are general, to determine the conditions for their use and to build up a knowledge base about AI techniques themselves. SUMEX resources are coordinated for this purpose with the multidisciplinary efforts of the Stanford Heuristic Programming Project (HPP). Under support from ARPA, NIH/NLM, ONR, NSF, and private funding, the HPP conducts research on five key scientific problem areas, as well as a host of subsidiary issues [1]:

- 1) Knowledge Representation - How shall the knowledge necessary for expert-level performance be represented for computer use? How can one achieve flexibility in adding and changing knowledge in the continuous development of a knowledge base? Are there uniform representations for the diverse kinds of specialized knowledge needed in all domains?
- 2) Knowledge Utilization - What designs are available for the inference procedure to be used by an expert system? How can the control structure be simple enough to be understandable and yet sophisticated enough for high performance? How can strategy knowledge be used effectively?
- 3) Knowledge Acquisition - How can the model of expertise in a field of work be systematically acquired for computer use? If it is true that the power of an expert system is primarily a function of the

quality and completeness of the knowledge base, then this is the critical "bottleneck" problem of expert systems research.

- 4) Explanation - How can the knowledge base and the line of reasoning used in solving a particular problem be explained to users? What constitutes an acceptable explanation for each class of users?
- 5) Tool Construction - What kinds of software packages can be constructed that will facilitate the implementation of expert systems, not only by the research community but also by various user communities?

Artificial Intelligence is largely an empirical science. We explore questions such as these by designing and building programs that incorporate plausible answers. Then we try to determine the strengths and weaknesses of the answers by experimenting with perturbations of the systems and extrapolations of them into new problem areas. The test of success in this endeavor is whether the next generation of system builders finds the questions relevant and the answers applicable to reduce the effort of building complex reasoning programs.

I.B Highlights

I.B.1 Handbook of Artificial Intelligence

The AI Handbook is a compendium of knowledge about the field of Artificial Intelligence being assembled under Professor Edward Feigenbaum and Messrs Avron Barr and Paul Cohen. It is being compiled by students and investigators at several research facilities across the nation. The AI Handbook Project is a good example of community collaboration using the SUMEX-AIM communication facilities to prepare, review, and disseminate this reference work on AI techniques. The Handbook articles exist as computer files at the SUMEX facility. All of our authors and reviewers have access to these files via the network facilities and use the document-editing and formatting programs available at SUMEX. This relatively small investment of resources has resulted in what we feel will be a seminal publication in the field of AI, of particular value to researchers, like those in the AIM community, who want quick access to AI ideas and techniques for application in other areas.

The AI Handbook Project was undertaken as a core activity by SUMEX in the spirit of community building that is the fundamental concern of the facility. We feel that the organization and propagation of this kind of information to the AIM community, as well as to other fields where AI is being applied, is a valuable service that we are uniquely qualified to support.

The scope of the work is broad: Two hundred articles cover all of the important ideas, techniques, and systems developed during 20 years of research in AI. Each article, roughly four pages long, is a description written for non-AI specialists and students of AI. Additional articles serve as overviews, which discuss the various approaches within a subfield, the issues, and the problems.

We expect the Handbook to reach a size of approximately 1000 pages. Roughly two-thirds of this material will constitute Volumes I and II of the Handbook. The material in Volumes I and II will cover AI research in Heuristic Search, Representation of Knowledge, AI Programming Languages, Natural Language Understanding, Speech Understanding, Automatic Programming, and Applications-oriented AI Research in Science, Mathematics, Medicine, and Education. Researchers at Stanford University, Rutgers University, SRI International, Xerox PARC, RAND Corporation, MIT, USC-ISI, Yale, and Carnegie-Mellon University have contributed material to the project. The current schedule for publication of the several volumes is as follows. It should be noted that Volume I has been selected by the Library of Computer Science as their August 1981 book club selection.

May, 1981: Publication of Volume 1 by publisher (Wm. Kaufmann Inc., Los Altos, Ca.)

August, 1981: Submission of final copy to publisher for Volume II (publication by end of 1981).

August-September, 1981: Completion of Technical Reports containing chapters of Handbook

October, 1981: Submission of final copy of Volume III to publisher (for publication first quarter 1982)

I.B.2 Tutorial on AI in Clinical Medicine

In conjunction with the 1980 AIM Workshop, a continuing education tutorial designed for physicians was held at Stanford on August 17-18, 1980. The tutorial was entitled "Computers in Medicine -- Applications of Artificial Intelligence Techniques" and was organized by Drs. W. Clancey and E. Shortliffe. The tutorial was well-attended by 135 physicians, 18 students, 10 members of the press, and several non-physician researchers. It was accredited for postgraduate medical education through Stanford University School of Medicine. Enrollees came from as far away as Mexico and the East Coast.

The course included an optional introduction to computers for those who had no prior experience with the technology, an overview of SUMEX-AIM research, and an introduction to background materials regarding decision theory and data base applications in medicine. Speakers also provided detailed presentations on MYCIN, CASNET/EXPERT, INTERNIST and GUIDON. The course closed with a panel discussion on the problems and promise of AI in Medicine. A syllabus was distributed including a comprehensive survey of medical AI research and is comprised of recent articles written by the tutorial faculty, mostly for a clinical audience. The faculty consisted of 15 distinguished researchers from the national AI community, including 7 physicians and 9 speakers from centers other than Stanford. Coordination and planning for the tutorial was facilitated by sending electronic messages; almost all speakers regularly use SUMEX or another ARPANET machine.

The course was exceedingly well received. Attendees were fascinated by the content, generally felt it was well presented, and indicated they would recommend the course to others if it were made available again. Many physicians requested a follow-up course that would introduce them to more technical detail than had been possible in the introductory tutorial.

To evaluate the impact of the tutorial on the participants, and to assess baseline opinions regarding the field, we undertook a survey of the physicians' knowledge about computers as well as their attitudes towards medical consultation systems. The statistical analysis of these questionnaires has now been completed, and a paper summarizing the results submitted for publication (*). In brief, the survey showed that physicians were willing to accept the possibility of computer-based clinical decision aids but placed severe demands on the capabilities of such systems if they were to be acceptable for routine use.

(*) Teach, R.L. and Shortliffe, E.H. "An Analysis of Physician Attitudes Regarding Computer-based Clinical Consultation Systems." Submitted for publication, March 1981.

I.B.3 GENET - Dissemination of AI Tools for Molecular Genetics

The MOLGEN project at Stanford has focused on applications of artificial intelligence and symbolic computation in the field of molecular biology. The research began in 1975 and by early 1980, through many collaborative contacts, it was realized that some of the systems developed by MOLGEN were already of direct utility to many scientists in the domain.

In order to broaden MOLGEN's base of scientist collaborators to molecular biologists at institutions other than Stanford and to experiment with the use of a SUMEX-like resource to disseminate sophisticated AI software tools to a generally computer-naive community, we initiated an experimental user group called GENET. The response to our very limited announcement of this facility has been most enthusiastic.

We have offered three main programs to assist molecular genetics users: SEQ, a DNA-RNA sequence analysis program; MAP, a program that assists in the construction of restriction maps from restriction enzyme digest data; and MAPPER (written and maintained by William Pearson from Johns Hopkins University), a simplified version of the MOLGEN MAP program that is somewhat more efficient than the MOLGEN version. Some of the other more sophisticated programs being developed by MOLGEN research efforts have not been offered because they are not ready for novice users. In addition, the GENET users have had access to the SUMEX-AIM programs for electronic messaging, text editing, file searching, etc.

The GENET community, begun in spring of 1980, started to grow exponentially until they were consuming SUMEX resources on a scale equal to the largest AI research project. We were obliged to place restrictions on the number of simultaneous GENET users and to otherwise limit the growth of the community. Even with these restrictions, the community currently consists of approximately 200 users from 63 research institutions. Of these 200 users, approximately 35 are consistently active users. That is, they log in, run programs, and interact with the MOLGEN members on an almost daily basis. Many of these users have made valuable contributions to our work. About 100 others are frequent, but not regular users. They log in only when they have a major analysis task to perform, which seems to be on the order of once a month.

The remaining users rarely use the system. They have logged in a few times, but for one reason or another they never become regular users of the system. Quite often this is because a lab group will settle on having one or two graduate students or post-doctoral associates become the "computer experts" of the group, and as a result, the computer use by the other people in the lab drops to a very low level. An equally prevalent reason for users to stop using the GENET account is a lack of SUMEX resources. The major complaint that we get from GENET users concerns the lack of computer time and availability of the system. One account just is not enough for 200 people to share.

We have succeeded in the goals set out for GENET. Many of our GENET guests have become active collaborators in core MOLGEN research. We are also pleased by the numerous comments SUMEX has received from GENET users

praising the user-sensitive nature of the resource, especially in comparison to typical university computer centers. It is clear we have only had the resources to whet the appetite of this large, active, international community.

I.B.4 AGE - A Tool for Knowledge-Based System Development

One of the most difficult, time-consuming, and expensive aspects of building knowledge-based systems (indeed any kind of software system) is the human effort involved in designing and coding them from the ground up. A major goal of SUMEX core research has been to demystify and make explicit the art of knowledge engineering. More concretely, we have attempted to isolate inference, control, and representation techniques from previous knowledge-based programs; reprogram them for domain independence; write an interface that will help a user understand what the package offers and how to use the modules; and make the package available to other members of the AIM community and the general scientific community to assist in knowledge-based program development. The AGE (Attempt to Generalize) package, developed by H. P. Nii and E. Feigenbaum is one of the earliest experimental examples of such a system and has reached a level of practical utility.

The design and implementation of the AGE program is based primarily on the experience gained with knowledge-based programs at the Stanford Heuristic Programming Project in the last decade. The programs that have been, or are being, built include: DENDRAL, meta-DENDRAL, MYCIN, HASP, AM, MOLGEN, CRYSLIS [Feigenbaum 1977], and SACON [Bennett 1978]. Initially, the AGE program embodies the AI techniques used in these programs but longer range goals are to integrate those developed at other AI laboratories as well. It is hoped that AGE will speed up the process of building knowledge-based programs and facilitate the dissemination of AI techniques by (1) packaging common AI software tools so that they need not be reprogrammed for every problem; and (2) helping people who are not knowledge engineering specialists write knowledge-based programs.

AGE is being developed along two separate fronts: the "kit" of tools for implementing knowledge-based systems and the "intelligent" interface to assist users make use of them. The current AGE system provides a set of preprogrammed "components" or "building blocks". A "component" is a collection of functions and variables that support conceptual entities in program form. For example, the production rule component, consists of a rule interpreter and various strategies for rule selection and execution.

The components in AGE have been carefully selected and modularly programmed to be useable in combinations. For those users not familiar enough to experiment on their own, AGE provides two predefined configurations of components--each configuration is called a "framework". One framework, called the Blackboard framework, is for building programs that use a globally accessible data structure called a "blackboard" [Lesser 77], and independent sources of knowledge which cooperate to form hypotheses. The other framework, called the Backchain framework, is for building programs that use backward-chained production rules as their primary mechanism of generating inferences (e.g., MYCIN).

Currently AGE-1 is available on a limited basis on the SUMEX-AIM resource and on the Stanford SCORE 2060 computer in the Computer Science Department. We held a three-day workshop in March 1980 to familiarize invitees with the use of AGE and to allow each participant to implement a

running program related to his application area. For the 1980 AIM Workshop we reimplemented a major portion of the VM program using AGE. In addition to demonstrating a variety of features of AGE, we were able to illustrate the relatively short implementation time required once the goals of the application and the necessary knowledge were delineated -- a first-year graduate student had the program running in three weeks. We are still working to broaden the user community for AGE and to learn from their experiences what directions our future research efforts should take.

I.B.5 ONCOCIN- An Oncology Chemotherapy Advisor

Work on the oncology chemotherapy consultation system, named ONCOCIN, was begun in July 1979. It is one of the newest application areas being investigated in the Stanford SUMEX-AIM community and is designed to be an interactive system for assigning and managing patients on chemotherapy protocols. This spring, it was installed for initial experimental use by faculty and fellows in the Debbie Probst Oncology Day Care Center at Stanford University Medical Center. Overall goals for ONCOCIN are (1) to demonstrate that a rule-based consultation system with explanation capabilities can be usefully applied and accepted in a busy clinical environment; (2) to improve the tools available for building knowledge-based expert systems for medical consultation; and (3) to establish an effective scientific relationship with a group of physicians that will facilitate future research and implementation of knowledge-based tools for clinical decision making.

In addition to ONCOCIN's basic AI research goals, it is directed toward the development of a clinically useful oncology consultation tool that will: (1) assist with the identification of protocols that may apply to a given patient and to help determine the patient's eligibility for a given protocol; (2) provide detailed information on protocols in response to questions from clinic personnel; (3) assist with chemotherapy dose selection and attenuation for a given patient; (5) provide reminders, at appropriate intervals, of follow-up tests and films required by the protocol in which a given patient is enrolled; and (6) reason about managing current patients in light of stored data from previous visits of the individual patients or aggregate data about groups of "similar" patients.

We are pursuing a five-year plan for accomplishing these goals. We spent the first year working out a prototype ONCOCIN system, drawing from programs and capabilities developed for the EMYCIN system-building project. We also undertook a detailed analysis of the day-to-day activities of the Stanford oncology clinic in order to determine how to introduce ONCOCIN with minimal disruption of an operation which is already running smoothly. Much of this early effort was spent giving careful consideration to the most appropriate mode of interaction with physicians in order to optimize the chances for ONCOCIN to become a useful and accepted tool in this specialized clinical environment.

More recently we have detailed the design and have implemented an actual experimental system. This system is based on multiple processes that manage the physician interface, reasoning and problem-solving, and patient database management. All of the system work has been completed to allow installation of ONCOCIN in the clinic. Following the initial prototype development based on lymphoma protocols, we checked to verify that the representation method we are using will be adequate for arbitrary protocol knowledge that may be encountered in the future. So we decided to encode and briefly test the knowledge of a non-lymphoma protocol. We chose the complicated protocol for oat cell (small cell) carcinoma of the lung because it involves a large number of possible therapies and complex interweaving of chemotherapy and radiotherapy. After approximately one

month's effort, the oat cell protocol was encoded and run successfully on a number of test cases. In addition, the lymphoma protocol specifications used in the clinic were changed and we spent a few weeks entering the necessary corrections. In all cases the ONCOCIN representation scheme was adequate to accommodate the protocol knowledge with only minor changes, and we are confident that the system will be able to adapt to other protocols that need to be encoded in the coming years.

ONCOCIN has been extensively debugged through runs on several hundred sample patient cases with the results reviewed in detail by the collaborating oncologists. We have just begun to offer the ONCOCIN system for use by the oncology faculty and fellows in the morning chemotherapy clinics in which most of the lymphoma patients receive their treatment. We have taken care in introducing ONCOCIN to provide needed baseline information so we can formally evaluate its impact and effectiveness in the oncology clinic.

I.C Administrative Changes

The SUMEX-AIM resource has undergone several administrative changes this past year that serve to enhance its position within the Stanford Medical School as a resource for AI research:

- 1) Professor Edward Shortliffe was appointed as co-Principal Investigator of SUMEX-AIM. Professor Shortliffe has been central in the development of the MYCIN group of projects and has long worked closely with Professor Feigenbaum in planning the future development of SUMEX. This appointment takes formal recognition of this role for Professor Shortliffe and strengthens SUMEX-AIM through his close scientific and administrative ties to the Stanford medical community.
- 2) In parallel with Professor Shortliffe's appointment as co-Principal Investigator, SUMEX moved administratively from the Department of Genetics to the Department of Medicine. It is now administered jointly between the Departments of Medicine and Computer Science. As part of the largest clinical medicine department at Stanford, SUMEX now has increased visibility and opportunity to broaden its local scientific collaborations.
- 3) Professor Elliott Levinthal began a two-year leave of absence to take a position as head of the Defense Sciences Office at DARPA. Professor Roy Maffly has replaced him as AIM liaison in charge of coordinating the reviews of new project applications and serving as the interface to collaborative projects.

I.D Resource Management and Allocation

The mission of SUMEX-AIM, locally and nationally, entails both the recruitment of appropriate research projects interested in medical AI applications and the catalysis of interactions among these groups and the broader medical community. User projects are separately funded and autonomous in their management. They are selected for access to SUMEX on the basis of their scientific and medical merits as well as their commitment to the community goals of SUMEX. Currently active projects span a broad range of application areas such as clinical diagnostic consultation, molecular biochemistry, belief systems modeling, mental function modeling, and instrument data interpretation (descriptions of the individual collaborative projects are in Section II beginning on page 89).

I.D.1 Management Committees

Since the SUMEX-AIM project is a multilateral undertaking by its very nature, we have created several management committees to assist in administering the various portions of the SUMEX resource. As defined in the SUMEX-AIM management plan adopted at the time the initial resource grant was awarded, the available facility capacity is allocated 40% to Stanford Medical School projects, 40% to national projects, and 20% to common system development and related functions. Within the Stanford aliquot, Prof. Feigenbaum and BRP have established an advisory committee to assist in selecting and allocating resources among projects appropriate to the SUMEX mission. The current membership of this committee is listed in Appendix C.

For the national community, two committees serve complementary functions. An Executive Committee oversees the operations of the resource as related to national users and makes the final decisions on authorizing admission for new projects and revalidating continued access for existing projects. It also establishes policies for resource allocation and approves plans for resource development and augmentation within the national portion of SUMEX (e.g., hardware upgrades, significant new development projects, etc.). The Executive Committee oversees the planning and implementation of the AIM Workshop series currently implemented under Prof. S. Amarel of Rutgers University and assures coordination with other AIM activities as well. The committee will play a key role in assessing the possible need for additional future AIM community computing resources and in deciding the optimal placement and management of such facilities. The current membership of the Executive committee is listed in Appendix C.

Reporting to the Executive Committee, an Advisory Group represents the interests of medical and computer science research relevant to AIM goals. The Advisory Group serves several functions in advising the Executive Committee; 1) recruiting appropriate medical/computer science projects, 2) reviewing and recommending priorities for allocation of resource capacity to specific projects based on scientific quality and

medical relevance, and 3) recommending policies and development goals for the resource. The current Advisory Group membership is given in Appendix C.

These committees have actively functioned in support of the resource. Except for the meetings held during the AIM workshops, the committees have "met" by messages, net-mail, and telephone conference owing to the size of the groups and to save the time and expense of personal travel to meet face to face. The telephone meetings, in conjunction with terminal access to related text materials, have served quite well in accomplishing the agenda business and facilitate greatly the arrangement of meetings. Other solicitations of advice requiring review of sizable written proposals are done by mail.

We will continue to work with the management committees to recruit the additional high quality projects which can be accommodated and to evolve resource allocation policies which appropriately reflect assigned priorities and project needs. We will continue to make information available about the various projects both inside and outside of the community and thereby promote the kinds of exchanges exemplified earlier and made possible by network facilities.

I.D.2 New Project Recruiting

The SUMEX-AIM resource has been announced through a variety of media as well as by correspondence, contacts of NIH-BRP with a variety of prospective grantees who use computers, and contacts by our own staff and committee members. The number of formal projects that have been admitted to SUMEX has more than trebled since the start of the project to a current total of 8 national AIM projects and 8 Stanford projects. Others are working tentatively as pilot projects or are under review.

We have prepared a variety of materials for the new user ranging from general information such as is contained in a SUMEX-AIM overview brochure to more detailed information and guidelines for determining whether a user project is appropriate for the SUMEX-AIM resource. A questionnaire is available to assist users seriously considering applying for access to SUMEX-AIM. Pilot project categories have been established both within the Stanford and national aliquots of the facility capacity to assist and encourage new projects in formulating possible AIM proposals and pending their application for funding support. Pilot projects are approved for access for limited periods of time after preliminary review by the Stanford or AIM Advisory Group as appropriate to the origin of the project.

These contacts have sometimes done much more than provide support for already formulated programs. For example, Prof. Feigenbaum's group at Stanford previously initiated a major collaborative effort with Dr. Osborn's group at the Institutes of Medical Sciences in San Francisco. This project in "Pulmonary Function Monitoring and Ventilator Management - PUFF/VM" (see Section II.A.2.4 on page 201) originated as a pilot